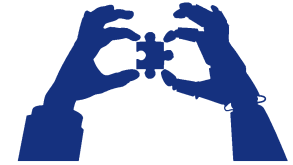


3

Human-Machine-Interaction(HMI)

- 사용자 중심의 지속가능한 인간 기계 상호작용



1

인간과 컴퓨터, 대화 형식의 상호작용

2022년 11월 30일에 출시된 ChatGPT는 인간 기계 상호작용에 있어 혁명을 불러왔다. 진정으로 인간이 기계와 대화로 상호작용할 수 있게 된 것이다. 물론 아직은 불완전하다. 그러나, 어쩌면 인류 역사상 최초로 인간 중심으로 기계와 상호작용하는 것이 가능하게 된 것이다. 인간은 기계, 즉 컴퓨터와 키보드로 상호작용해왔고, 지금도 그러하다. 키보드는 인간보다는 기계를 위한 인터페이스다. 인간이 키보드에 익숙해져야만 사용할 수 있는 것이다. 이후 마우스가 나왔고, GUI(Graphic User Interface)로 컴퓨터와 인간이 상호작용하게 되었다. 파일을 지우기 위해서는 마우스로 파일의 아이콘을 끌어다가 휴지통 모양의 아이콘으로 보내면 지울 수 있게 되었다. 그러면, 휴지통이 불룩해졌다. 지운

이경전

경희대 교수. 빅데이터 연구센터 소장. AI & BM Lab지도교수.

사용자중심인공지능포럼(UCAIFORUM.org) 연구원장. 한국경영정보학회 부회장 및 AI 연구회 회장. 미국인공지능학회(AAAI: American Association for AI) 혁신적인인공지능 응용(Innovative Applications of AI) 1995, 1999년, 2020년 수상. 2018년 전자정부유공자 대통령표창. 국제전자상거래연구원장, 한국지능정보시스템학회 회장(2017) 역임. 2021-2023년 한국시스타트업 100 선정위원장. 2022년 3월, 한국공학한림원 회원으로 선출. KAIST 경영과학 학·석·박사 졸업, 서울대 행정학 석·박사를 수료했다. CMU 로보틱스 연구소 초빙과학자(1996-97), MIT(2009-10)와 UC버클리(2010) Fulbright 초빙교수로 연구하였으며, AI Magazine에 논문 세 편을 게재하는 등 국제학술지에 40여 편의 논문을 발표했다. 인공지능과 디지털 비즈니스 모델 등을 연구한다(한국연구재단 중점 연구소).

파일을 다시 살리고 싶으면, 블록해진 휴지통 아이콘을 마우스로 클릭하여 다시 살릴 수 있었고, 그러면 블록했던 휴지통은 다시 훌쩍해졌다. 애플의 매킨토시가 처음으로 그러한 모습을 선보였다. 인간은 그림(그래픽)으로 컴퓨터와 상호작용할 수 있게 되었다.

그래픽 사용자 인터페이스에 터치스크린이 응용되면서, 인간은 기계의 화면을 직접 만지면서 상호작용하게 되었다. 이른바 Touch 인터페이스가 시작되었다. 전조는 여러 번 있었으나, 애플의 iPhone이 이를 완전히 대중화시켰다. 현대인들은 스마트폰의 화면을 너무 많이 터치하는 바람에 지문이 닳아 없어질 지경이다. 이제는 스마트폰 사용을 위한 끝무까지 나오고 있는 상황이다.

터치스크린에 기반한 인간 기계 상호작용에 이어 다음의 상호작용 기술은 무엇일까? NFC(Near Field Communication)기술도 나왔고, BLE(Bluetooth for Low Energy)기술도 나왔다. 그러나, 키보드-마우스-터치스크린으로 이어지는 주류적 변화를 가지고 오지는 못했다. 컴퓨터와 대화(Conversation)로 상호작용한다는 아이디어는 오랜 기간 동안 시도되어 왔고, 영화 HER나 아이언맨의 J.A.R.V.I.S.로 예고되어왔으나 좀처럼 실현되지 못하고 있었다. 스티브 잡스는 애플의 iPhone 4S에 시리를 도입하였으나, 사람들은 대화로 아이폰과 상호작용하는 것으로 옮겨가지 못했다. 구글 역시 안드로이드에 어시스턴트를 도입했으나, 시리와 마찬가지로의 결과를 가져왔다. 그런 과정에서 OpenAI의 ChatGPT가 이 모든 것을 바꿔놓게 되었다.

기계와 대화로 상호작용하려면 기계가 인간의 대화의 문맥을 이해하여야 한다. 문맥을 어떻게 이해할 수 있을까? 현재는 Transformer(Vaswani et al. 2017) 라는 기술이 그것을 어느 정도 가능하게 해주고 있다. Transformer 모델은 Attention(주의집중)이라는 메커니즘(Bahdanau, Cho, & Bengio 2016)을 도입한 인공지능망 모델이다. 대화와 같이, 순서가 있는 데이터를 기반으로 하는 인공지능망 모델은, 트랜스포머 모델의 발명 이전에는 RNN(Recurrent Neural Network)이라는 기법으로 모델링 되어왔는데, 2017년 구글의 연구자들이 발표한 트랜스포머는 어텐션 스코어를 한꺼번에 계산하는 방식으로 하여, RNN보다 계산의 효율성을 극적으로 향상시켰다. 학습 초기에는 예측 결과가 좋지 않았지만, 학습을 반복해 나감에 따라 가장 적절해 보이는 예측 결과를 출력하도록 계속 인공지능망의 파라미터들을 학습시키는 모델이다.

OpenAI의 ChatGPT는 구글이 발표한 트랜스포머 모델에서 디코더 부분만을 취한 반쪽 모델을 GPT(Generative Pre-Trained Transformer)라는 이름을 붙여 GPT-1(Radford et al. 2018), GPT-2(Radford et al. 2019), GPT-3(Brown et al. 2020), GPT-3.5(2022), GPT-4(2023)으로 발전시켜왔으며, GPT 3.5는 파라미터의 수가 1750억 개, GPT-4는 1조760억 개로 추정되고 있는 초거대 인공지능망이다.

GPT-4는 32768개의 토큰(영어 단어로 약 25000단어, 한국어 단어로 약 20000단어)을 입력으로 받을 수 있는 기계로, 인간이 GPT-4라는 기계와 상호작용하는 것은 한국인의 경우 약 20000단어를 입력

으로 줄 수 있는 것을 의미한다. 즉, 인간은 이제 기계와 수만 단어로 상호작용할 수 있게 되었다. 이것은 매우 당황스러운 환경의 변화다. 갑자기 기계와 대화가 가능해지게 되었는데, 대화가 가능해지자마자 기계와 수만 단어로 대화할 수 있게 되어 오히려 인간이 자신의 능력에 좌절감을 느끼는 상황이 갑자기 되어 버린 것이다.

인간들은 이러한 상황에 당황하지 않고, 대응하고 있는데, 그것이 바로 프롬프트 엔지니어링이다. 영어 기준으로 인간에게 25000단어로 이루어진 질문을 할 수 있는 상황에서, 인간들은 인간 세계에서 나눌 수 있는 대화의 수준에다가 자동으로 질문을 보완하는 단어들과 명령어를 첨가함으로써, 컴퓨터에 더욱 정교한 질문과 명령을 하는 것을 시도하고 있다. 그리고, 컴퓨터가 대답한 내용을 자동으로 가공하여 인간에게 제공하는 형태로 서비스가 진화하고 있다.

즉, 인간과 컴퓨터의 상호작용이, 1) 인간과 컴퓨터의 상호작용, 그리고 2) 컴퓨터와 인공지능의 상호작용으로 나누어져서 계층적으로 진행되고 있는 것인데, 이러한 계층화는 점점 더 심화될 것이다. 즉, 인간과 컴퓨터 간의 상호작용을 위해, 그 중간에 컴퓨터와 컴퓨터 간의 복잡한 상호작용이 내재되게 될 것이라는 점이다. 그런데, 이 모든 것이 인간의 대화를 닮은 모습으로 진행된다는 것이 OpenAI의 ChatGPT가 가져온 혁명이다.

인간과 컴퓨터의 대화, 프라이버시가 필요

우리 인간과 인간의 대화를 상상해보자. 오늘 당신이 다른 사람들과 나눈 대화는 모두 공적인가? 다른 사람에게 공개되어도 되는 대화를 했는가? 그렇지 않을 것이다. 그렇다. 많은 대화는 사적이다. 우리가 나눈 대화가 언젠가는 다 공개되어야 할 내용이라면, 우리는 좀처럼 대화를 하지 않고 입을 다물게 될 것이다. 인간과 컴퓨터 간의 대화도 마찬가지다. 인간과 컴퓨터 간의 대화가 사적이라면, 그 대화 상대인 컴퓨터가 저 멀리 OpenAI나 Naver 클라우드 같은 서버에 있는 인공지능 컴퓨터가 아니라, 내 스마트폰이나 랩톱 컴퓨터에 설치되어 있는 인공지능 시스템이어야 하는데, 이것이 가능한가? 결론부터 말한다면 불가능한 일이다.

사실 현재 ChatGPT, GPT-4 수준의 인공지능 시스템은 몇 년 이내에 충분히 우리의 컴퓨터에 설치될 수 있는 수준의 것이다. 현재 OpenAI는 그러한 비즈니스 모델을 추구하고 있지 않다. 그러나, 메타의 LLaMA(Touvron et al. 2023), 이를 활용한 스탠포드대의 알파카(Taori et al. 2023), 비쿠나(Chiang et al. 2023), Stability.AI의 StableLM이나 MosaicML의 MPT 등 이른바 s-LLM(Small Large Language Model)은 일반 사용자의 컴퓨터나 스마트폰에 설치하여 사용하는 인공지능 비즈니스 모델을 추구하고 있다. 현재 OpenAI

의 GPT 시리즈나 네이버의 하이퍼클로바 시리즈와 같은 Closed Source Large Language Model이 몇 년 후에 설치형으로 콤팩트화 될 수도 있고, 라마나 스테이블LM, MPT 등의 OpenSource s-LLM이 충분한 성능을 보여주면서 설치형으로 정착할 수도 있다. 어떤 경우든 우리는 몇 년 안에 내 랩톱 컴퓨터나 스마트폰에 장착된 인공지능과 사적인 대화를 하는 행복을 누릴 수 있게 될 것이다. 이 경우 모든 서비스의 형태는 에이전트 기반의 서비스가 될 것이다. 즉, 인간은 자신의 인공지능 에이전트와 사적인 대화를 통한 상호작용을 하고, 그 인공지능 에이전트는, 주인님인 나 대신에 여러 다른 인공지능 에이전트와 사적, 공적인 대화를 통한 상호작용을 하게 될 것이다.

3

연합학습 기술 사용

이러한 상황에서 각 인공지능 에이전트는 어떻게 자신의 성능을 높일 것인가? 그것을 가능하게 하는 기술이 연합학습(Federated Learning)이다. 모든 개인은 사생활 보호를 위해 로컬에 있는 인공지능 에이전트와 대화한다. 이 과정에서 대화는 이 에이전트들의 지능 향상을 위해 학습데이터로 활용될 수 있으나, 그 대화 데이터가 한곳에 모여서 저장되면 프라이버시 침해와 데이터 자산의 공개, 고객 정보의 유출 등 여러 문제가 발생할 수 있

다. 사용자와 인공지능 에이전트 사이의 대화는 로컬 컴퓨터를 떠나지 않으면서, 인공지능 에이전트 공급사 또는 플랫폼사가 AI 모델을 사용자 로컬 컴퓨터에 보내서 로컬에서 학습을 시킨 후 그 학습된 파라미터를 가중 평균하여 글로벌한 인공지능 모델을 구축하는, 이른바 개인화된 연합 학습(Personalized Federated Learning) 형태로 진행할 수 있다(Tan et al. 2022).

이렇게 할 경우 사용자는 자신의 대화 데이터를 노출하지 않을 수 있게 되므로, 사용자의 프라이버시, 보안성이 강화되고, 자신의 데이터를 플랫폼에 빼앗기지 않을 수 있으므로 사용자와 플랫폼 간의 데이터 소유 불균형이 더 심해지지 않으므로 생태계의 지속 가능성과 강건성을 높이게 될 것이다. 또한 이렇게 되면 사용자들은 자기의 로컬의 데이터를 더욱 소중하게 가꾸는 일을 하게 될 것이다. 이른바 모델 중심의 AI가 아닌 데이터 중심(Data-Centric)의 AI를 하게 되며, 데이터 부족과 AI 전문가 부족에 시달리는 개인이나 소상공인, 개인사업자, 중소기업에게는 이러한 방식이 희소식이 될 것이다. 즉, 연합 학습 기술은 디지털 디바이드, AI 디바이드, 데이터 갭을 줄여서 디지털 AI 생태계에 지속가능성을 높이는데 기여한다. 또한 개인은 자신의 데이터에 특별히 FineTuning된 로컬 AI 모델을 가질 수도 있게 되므로, AI 모델이 획일적이지 않게 되어 AI 모델의 다양성을 생태계가 가지게 되고, 앞서 설명한 것처럼, 개인과 플랫폼 간의 데이터 보유에 균형이 생겨서, 개인을 Empowering하는 결과에 기여하게 된다. 이른바, 인간 중심의 인공지능, 사용자 중심의 인공지

능(UCAI: User-Centric AI)이 되는 것이다. 이렇게 LLM을 연합학습과 결합하는 시도는 중국 텐센트(Tencent)의 위뱅크(WeBank)가 계속 개발해온 연합학습 프레임워크인 FATE에 LLM을 결합하는 시도가 있고, DynamoFL, FedML 등의 회사가 있으며, 사용자 중심 인공지능 포럼(UCAIFORUM.org)에서 연구하고 있다.

연합학습 방법론은 인간과 컴퓨터 간의 상호작용 데이터에도 활용되지만, 인간이 생성하는 모든 데이터에도 적용되어야 할 것이다. 인간의 몸은 순간순간 계속 데이터를 산출하고 있다. 독점 플랫폼 중심의 사고 방식에서는 이러한 모든 인간 데이터가 플랫폼에 저장되는 것을 가정한다. 그러나, 이러한 것은 인간 중심, 사용자 중심의 방법이 아니라 독점 플랫폼 중심의 방법이다. 대부분의 헬스케어 서비스 플랫폼도 사용자, 환자, 고객의 모든 생체 데이터를 자신의 서버에서 관리한다는 패러다임에서 벗어나지 못하고 있다. 이것은 연합학습 방법론이 2016년 구글에 의해서 지원되고 개발(Konečný et al. 2016; McMahan et al. 2017)되기 이전이라면 어쩔 수 없는 일이다. 그러나, 이제 이미 연합학습 방법론이 나왔고, 그 응용이 계속 Nature 등의 세계 최고 학술지에 보고되고 있는 현실을 볼 때(Bai et al. 2021), 독점 플랫폼 중심의 방법론은 좀 더 인간 중심, 사용자 중심의 지속가능성을 높이는 민주적 플랫폼 방법론에 의해 대체될 것이다.

개인이 컴퓨터와 많은 대화를 나눈 데이터이든, 개인의 생체 데이터이든, 개인의 행동 데이터이든 이 모든 개인 데이터들은 이제 개인의 소유와 관리에 둘 수 있어야 한다. 그것이 인간 중

심, 사용자 중심의 인간 기계 상호작용일 것이다.

4

대화 → 거래, 거래 → 자동 행동 - 인간 기계 상호 작용이 에이전트화(化)

ChatGPT로 대표되는 현재의 초거대 인공지능 서비스는 주로 인간과 컴퓨터 간의 대화로 이루어지지만, 점차 대화는 거래로 변화될 것이다. 아마존, 이베이, 인터파크, 쿠팡 등으로 대표되는 기존의 전자상거래는 대화형 전자상거래로 변모하여, 이제는 인터넷이 아닌 인공지능 그 자체가 거래를 촉진하게 될 것이다. 그런데, 그러한 시간도 잠깐뿐. 인간은 컴퓨터, 즉, 기계에게 자동 행동을 요구할 것이다. 그것이 오랫동안 논의되어 온 에이전트화이다(이경전 1999, Jin, Suh, & Lee 2003; Jin & Lee 2001; Lee, Chang, & Lee 2000). 인간은 인공지능 에이전트에게 대화로 명령한다. 그런데, 어떤 일의 단위에 관한 명령을 주는 것이 아니라, 큰 목표를 줄 수 있다. 인간으로부터 목표를 부여받은 인공지능 에이전트는 자신의 목표를 하위목표로 나누고, 이 하위 목표 각각을 달성하는 일의 순서를 계획하고, 이를 최적의 자원 활용과 순서로 실행한다. 이러한 미래를 지금 보여주는 것들이, AutoGPT, AgentGPT와 같은 사례다.

마이크로소프트 창업자 빌 게이츠는 미래의 최고 기업은 AI를

활용한 Personal Digital Agent를 만드는 회사가 될 것으로 예측하면서[•], Personal Digital Agent는 사람을 대신해 특정 일을 수행할 수 있는 AI 개인비서를 가리키는데, AI가 개인비서 형태로 발전한다면 구글 같은 검색 사이트와 아마존 같은 사이트를 번거롭게 직접 방문할 필요 없이 AI Agent가 모든 것을 알아서 처리할 것으로 예상하기도 했다. 빌 게이츠는 Inflection AI라는 회사를 주목하고 있다고 했는데, 이 회사는 AI 개인 비서를 개발 중인 스타트업 구글 답마인드 임원 출신의 무스타파 슬레이먼과 링크드인 설립자 리드 호프먼이 공동 창업한 스타트업으로 대화형 개인비서인 Pi(파이)를 2023년 5월에 출시하였다. Pi는 Inflection AI가 자체 개발한 LLM을 기반으로 만들어졌으며 단순히 질문에 대답하는 것이 아닌 개인적인 문제에 대한 대화와 조언 등 인간과의 상호작용을 위해 만들어진 생성 AI로 Inflection AI는 현재 세계 정보에 접근할 수 있는 방법이 될 생성 AI를 개발하고 있으며 개인의 온라인 구매, 예약 같은 일을 돕는 데 최적화된 AI 비서를 만드는 것이 목표라고 밝히고 있다.

인간은 삶의 의미를 찾으며 동시에 즐거움을 추구하며 살고 있다. 인간은 다른 동물과 달리 도구를 사용하며, 지금까지 인간이 개발한 가장 강력한 도구는 기계, 특히 컴퓨터와 연결된 기계이고, 인간과 대화로 작동할 수 있는 기계이다. 이러한 기계와 상호작용할 때, 인간 중심으로, 유연하면서도, 어디에 치우침이 없

이 지속 가능한 상호작용을 한다는 것은 이상(Ideal)과도 같은 것이다. 그러한 이상에 도달하기 위해 인류는 끊임없이 노력해왔고, 최근 ChatGPT 등 생성형 AI, 대화형 AI의 발전은 인류의 이상에 갑자기 몇 발자국을 앞당기는 큰 성과이다. 이러한 기술은 이제 연합 학습과 같은 인간 사용자와 조직의 데이터를 보호하는 동시에 인공지능 성과를 높이는 방법론과 결합하여 인간 중심의 기술로 발전할 수 있게 되었다. 그리고 이렇게 발전된 기술은, 인간이 많은 일들을 기계라는 에이전트에 맡길 수 있는 수준을 실현할 수 있게 하고 있다.

인간에게는 새로운 능력이 요구되고, 새로운 임무가 부여된다. 인공지능 기계와 더욱 전문적으로 커뮤니케이션하면서 일의 성과를 높이는 능력을 갖출 필요가 생긴다. 예를 들어, GPT-4에게 20000단어로 이루어진 명령을 효율적, 효과적으로 내릴 수 있는 능력이 필요한 것이다. 그리고, 그렇게 빠르게 일하고, 콘텐츠를 생성하는 AI의 결과를 잘 평가할 수 있는 인간으로서의 지식과 능력, 지혜, 감각이 필요한 것이다. 이러한 능력은 이제 인간의 임무와 책임으로 연결될 것이다. 근미래에 인간은 쉴 때, 인공지능에게 일을 시켜야 하는 공식적, 비공식적 책임과 임무를 부여받게 될 것이다. 인공지능을 잘 사용하는 능력을 갖추어야 하며, 인공지능에게 일을 잘 시키는 새로운 책임을 부여받게 되는 것이다.

• <https://www.aitimes.com/news/articleView.html?idxno=151273>, <https://www.goldmansachs.com/intelligence/pages/bill-gates-on-the-ai-revolution.html>

참고 문헌

- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention Is All You Need. arXiv:1706.03762
- Bahdanau, D., Cho, K., & Bengio, Y. (2016). Neural Machine Translation by Jointly Learning to Align and Translate. arXiv:1409.0473.
- Radford, A., Narasimhan, K., Salimans, T. & Sutskever, I. (2018). Improving language understanding by generative pre-training. <https://openai.com/research/language-unsupervised>
- Radford, A., Wu, J., Child, R., Luan, D., Amodei, D. & Sutskever, I. (2019). Language models are unsupervised multitask learners, https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_learners.pdf
- Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., Hesse, C., Chen, M., Sigler, E., Litwin, M., Gray, S., Chess, B., Clark, J., Berner, C., McCandlish, S., Radford, A., Sutskever, I., & Amodei, D. (2020). Language Models are Few-Shot Learners. arXiv:2005.14165
- Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M-A., Lacroix, T., Rozière, B., Goyal, N., Hambro, E., Azhar, F., Rodriguez, A., Joulin, A., Grave, E., & Lample, G. (2023). LLaMA: Open and Efficient Foundation Language Models. arXiv:2302.13971
- Taori, R., Gulrajani, I., Zhang, T., Dubois, Y., Li, X., Guestrin, C., Liang, P., & Hashimoto, T. B. (2023). Alpaca: A Strong, Replicable Instruction-Following Model, <https://crfm.stanford.edu/2023/03/13/alpaca.html>
- Chiang, W.-L., Li, Z., Lin, Z., Sheng, Y., Wu, Z., Zhang, H., Zheng, L., Zhuang, S., Zhuang, Y., Gonzalez, J. E., Stoica, I., & Xing, E. P. (2023, March). Vicuna: An Open-Source Chatbot Impressing GPT-4 with 90%* ChatGPT Quality. <https://lmsys.org/blog/2023-03-30-vicuna/>
- McMahan, H. B., Moore, E., Ramage, D., Hampson, S., & Agüera y Arcas, B. (2017). Communication-Efficient Learning of Deep Networks from Decentralized Data. Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS) 2017. JMLR: W&CP
- Konečný, J., McMahan, H. B., Yu, F. X., Richtárik, P., Suresh, A. T., & Bacon, D. (2016). Federated Learning: Strategies for Improving Communication Efficiency. arXiv:1610.05492
- Tan, A. Z., Yu, H., Cui, L., & Yang, Q. (2022). Towards Personalized Federated Learning. IEEE Transactions on Neural Networks and Learning Systems, pp. 1-17. <https://doi.org/10.1109/tnnls.2022.3160699>
- Bai, X., Wang, H., Ma, L., Xu, Y., Gan, J., Fan, Z., Yang, F., Ma, K., Yang, J., Bai, S., Shu, C., Zou, X., Huang, R., Zhang, C., Liu, X., Tu, D., Xu, C., Zhang, W., Wang, X., ... Xia, T. (2021). Advancing COVID-19 diagnosis with privacy-preserving collaboration in artificial intelligence. Nature Machine Intelligence, 3, 1081-1089.
- 이경진, “전자상거래를 위한 소프트웨어 에이전트”, 정보처리학회지, 6(1):54-62, 1999.
- Jin, D., Suh, Y., Lee, K., “Generation of Hypotheses on the Evolution of Agent-Based Business Using Inductive Learning,” Electronic Markets, vol. 13, no. 1, 13-20, 2003.
- Lee, K. J., Chang, Y. S., Lee, J. K., “Time-Bounded Negotiation Framework for Electronic Commerce Agents”, Decision Support Systems, vol. 28, no.4, pp. 319-331, June, 2000.
- Jin, D. Lee, K, “Impacts and Limitations of Intelligent Agents to Internet Commerce,” Lecture Notes in Computer Science, vol. 2105, pp. 33-4